THE VALUE OF INDIRECT SELECTION
II. PROGENY TESTING

BU-585-M June, 1976

S. R. Searle
Biometrics Unit, Cornell University, Ithaca, New York

Abstract

Conditions under which progeny testing gives more rapid genetic improvement than performance testing are discussed in terms of genetic parameters. So also are conditions under which progeny testing using an alternative trait is preferred – i.e., indirect selection. Procedures for comparing alternative traits are also given.

This is the companion paper to that of Biometrics 21, 682-707, 1965, in which similar conditions are presented for mass selection.

1. Introduction

The efficiency of indirect selection relative to direct selection, under conditions of mass selection, is considered at length in Searle [1965]. Applications of results given there have since been seen in studies of egg production, body weight and other traits of the domestic fowl (e.g., Nordskog et al. [1967] and Festing and Nordskog [1967]) and in studies of milk yield and other traits in dairy cows (e.g., Eisen [1966], Thompson and Loganathan [1968], and Wilton and Van Vleck [1968]). And use has also been made of the relative selection index in fertility studies of Romney ewes (Ch'ang and Rae [1972]) and in at least one simulation study (Singh et al. [1967]). All these uses have been under circumstances

---

Paper No. BU-585-M in the Biometrics Unit Mimeograph Series, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853.

of mass selection. This paper (the belated companion to the first of this series of two) describes properties of relative selection efficiency under conditions of progeny-testing.

2. Relative Selection Efficiency of Progeny-Testing

A reasonably comprehensive discussion of improving a trait (hereafter referred to as the basic trait) through selection on other, alternative traits is given in Searle [1965]. We therefore go to the heart of the matter here, insofar as progeny-testing is concerned, with a minimal description of the notation used.

We consider the improvement, through selection, of a basic trait with genotype represented by y and phenotype by Y. Selection on Y is called direct selection. Selection on some index I is called indirect selection. We confine attention to just the additive portion of the genotype, and accordingly define heritability in the narrow sense as $h = \sigma_y^2/\sigma_Y^2$, the ratio of the additive genetic variance σ_y^2 to the phenotypic variance σ_Y^2. If R_{Iy} is the correlation between I and y then the relative selection efficiency of indirect selection on I compared to direct selection on Y is

$$RSE(I, Y, y) = R_{Iy}/\sqrt{h} \ . \qquad (1)$$

This is on the basis of assuming (as we will) that the intensity of selection is the same using I as it is using Y. Before introducing selection on traits other than Y we look first at the efficiency of the progeny-test selection index itself.

The circumstances we deal with are those where the selection index is based on phenotypes of progeny of the animals (i.e., parents) among whom selection is to be practiced. Two commonly occurring cases are those of using milk yields of dairy cows to practice selection among bulls (particularly for bulls used in artificial insemination programs); and of using egg production of domestic fowl

to practice selection among cockerels (their sires). It will be assumed, for simplicity's sake, that each sire (or, more generally, parent) being tested has the same number of progeny, n, and that each progeny has exactly one record. Although this sounds restrictive it imposes only minimum limitations in many practical applications; e.g., selection practiced on dairy and poultry sires is largely based upon just first records of their progeny. Provided there are sufficient progeny with records available at the end of the first possible year of such availability (e.g., approximately 60-100 progeny), then it is often more economic to make selections then (and bring in more young untried sires for testing) than it is to keep all sires additional years while waiting for repeat records on their progeny. We therefore assume that each sire has n progeny each with one record.

Suppose that the index I_Y is the selection index based on n progeny records on Y. Then it is well-known (e.g., Falconer [1960]) that $R(I_Y, y)$ for (1) is

$$R(I_Y, y) = \sqrt{\frac{nh}{4 + h(n - 1)}} \qquad (2)$$

and hence

$$RSE(I_Y, Y, y) = \sqrt{\frac{n}{4 + h(n - 1)}} \; . \qquad (3)$$

Equation (2) gives the correlation between the genotype y and the progeny-test selection index based on n progeny each with one record Y. Values of this correlation are shown in Table 1 for numbers of progeny ranging from 2 to 200 and

(SHOW TABLE 1)

for heritability values from 0.1 to 1.0. A discussion of this table in the context of dairy cow breeding is given in Searle [1964], including reference to a

study of observed progeny tests that confirms the underlying derivation of this correlation.

Equation (2) gives opportunity to calculate the number of progeny required in order that the correlation between the index and the (additive part) of the sire's genotype shall be of a predetermined magnitude. Thus on rewriting (2) as

$$n = \frac{4/h - 1}{[R(I_Y, y)]^{-2} - 1} \qquad (4)$$

we can calculate n for any value of h and for a pre-assigned value of $R(I_Y, y)$. Table 2 (taken from Searle [1964]) shows values of n for the same ranges of $R(I_Y, y)$ and h as are used in Table 1. Naturally, the numbers of progeny required

(SHOW TABLE 2)

are small for correlations so low as to be valueless (below 0.80 say) but thereafter the numbers increase rapidly. Tables 1 and 2 are, of course, two different ways of tabulating values of the same expression and so their entries correspond. Nevertheless, the tables represent two different uses for the same result: Table 1 provides the correlation between additive genotype and progeny-test for some known number of progeny, whereas Table 2 is useful for deciding, prior to conducting a progeny-test program, just how many progeny are needed in order to achieve a pre-assigned value of the correlation $R(I_Y, y)$.

The poultry and dairy cow examples of progeny testing cited earlier are cases where the trait of interest (egg-laying and milk production) cannot be expressed by the sire. Consequently selection on sires could never be based on performance records and so the RSE of equation (3) is not applicable. Progeny-testing is usually the preferred method of estimating additive genotype in such cases,

although an alternative is to use ancestor records, the efficiency of which is discussed in Searle [1963]. But there are economically important traits which are expressable in both parent and offspring, such as fleece weight in sheep and rate of weight gain in beef cattle. In these cases the RSE of (3), which compares progeny testing with performance testing, is applicable and can be used to answer the question "when is progeny-testing preferred over performance testing?". By "preferred over" or "better than" is meant "gives greater rate of genetic improvement". And in this sense progeny-testing will be preferred over performance testing when RSE of (3) exceeds unity.

Values of (3) can be studied in two ways just as are the values of (2) in Tables 1 and 2. First, in Table 3, are shown values of RSE given by (3) for the

(SHOW TABLE 3)

same values of n and h as in Tables 1 and 2. These values are, by nature of (2) and (3), just those of Table 1 multiplied by $h^{-\frac{1}{2}}$. Nevertheless they are illuminating: for example, an immediate conclusion is that RSE > 1 on most occasions, i.e., except for very small n or for large h, progeny-testing is preferable to performance testing. Another feature of Table 3 is that for each value of h there is an upper limit on the value of the RSE. This is obtained by letting n tend to infinity in (3) giving $RSE(n \to \infty) = h^{-\frac{1}{2}}$, the values of which are shown in the last row of Table 3. For example, with an infinitely large number of progeny the relative selection efficiency of progeny-testing over performance testing is 3.16, for h = 0.1, but only 1.58 for h = 0.4. In making these observations it must not be forgotten that we are assuming single records on all animals. Were there to be m records in the performance test of the parent the RSE of (3), i.e., the values in Table 3, would be lessened through multiplication by $\sqrt{\rho + (1 - \rho)/m}$ where ρ is repeatability.

The second way to look at values of n, h and RSE that satisfy (3) is that of asking the same kind of question that leads to Table 2: in this case, "for a given value of h, how many animals are needed in a progeny-test so that the RSE equals some pre-assigned value?" This results in rearranging (3) as

$$n = \frac{4 - h}{(RSE)^{-2} - h} . \qquad (5)$$

Values of this expression are shown in Table 4, for the same values of h as in

(SHOW TABLE 4)

Tables 1 - 3, and for a range of values of RSE. The latter are chosen as equal or greater than unity, representing situations when progeny-testing is as good as or better than performance testing. The upper limit of $h^{-\frac{1}{2}}$ on RSE established from (3) as $n \to \infty$ and shown in Table 3 is clearly evident in (5), arising from the non-negativity of n. This is reflected in Table 4 by each column ending with its upper limit of RSE shown in parentheses, indicating that no further values of n exist. The choice of RSE values shown in the table was made with these upper limits in mind, so as to have a range of values as well as values that are, in most cases, close to the limits. As with the correspondence between Tables 1 and 2, so also is there a correspondence between the entries in Tables 3 and 4. But again the two tables have separate uses. Table 3 gives the RSE for a given number of progeny and a given heritability, whereas Table 4 is useful for deciding, prior to conducting a progeny-test program, just how many progeny are needed in order to achieve a pre-assigned increased rate of genetic improvement over performance testing, i.e., a pre-assigned value of RSE. It is readily seen that for low values of heritability just a relatively small number of progeny (e.g., 10 or less) can give RSE values in the neighborhood of 1.15 (up to 1.40 for h = 0.1); and that for high

heritability values (0.7 and greater) progeny-testing can never be better than having an RSE of 1.19. In contrast, for low heritability values but several hundreds of progeny it is possible to have RSE's in the neighborhood of 2.00, i.e., in the sense being used here, progeny-testing is twice "as good as" performance testing. All this is, of course, what one would expect, but it is interesting to have comparisons available, as provided by the tabulated values.

3. Progeny-Testing With An Alternative Trait

Taking I as the phenotype X of some trait alternative to Y reduces (1) to $r\sqrt{h_x/h_y}$ which we call

$$p = \text{RSE}(X, Y, y) = r\sqrt{h_x/h_y} \quad , \tag{6}$$

where r is the genotypic correlation between traits X and Y and $h_x = \sigma_x^2/\sigma_X^2$ is the heritability of X in the same way that $h_y = h = \sigma_y^2/\sigma_Y^2$ is of Y. Considerable attention was given to p in Searle [1965], and also to R(I, Y, y) for I being a linear combination, under mass selection, of either Y and one X, or of two X's. We now consider R(I, Y, y) for I being a progeny-test index.

3.1. Comparisons with performance testing

Denote by I_X the selection index based on single records of n_x progeny. Then, just like (2)

$$R(I_X, y) = r\sqrt{\frac{n_x h_x}{4 + h_x(n_x - 1)}} \tag{7}$$

and similar to (3)

$$\text{RSE}(I_X, Y, y) = r\sqrt{\frac{h_x}{h_y}}\sqrt{\frac{n_x}{4 + h_x(n_x - 1)}} = p\sqrt{\frac{n_x}{4 + h_x(n_x - 1)}} \quad , \text{ using (6).} \tag{8}$$

Expressions (7) and (8) are respectively the correlation between progeny-test on X and additive genotype y of a parent; and the relative selection efficiency for improvement in y, of progeny-testing using X compared to performance testing using Y. They are therefore directly comparable with (2) and (3). Furthermore, they are easily obtained from (2) and (3).

First, from (2) and (7) we see that (7) is (2) multiplied by r and using n_x and h_x for n and h; i.e.,

$$R(I_X, y) = r[R(I_X, x)].$$

Hence values of $R(I_X, y)$ from (7) are obtainable from Table 1 by multiplication by r, the genetic correlation. For example, for $r = 0.8$, $h_x = 0.7$ and $n_x = 50$, we get $R(I_X, y) = 0.8(0.95) = 0.76$. Also, values of n_x obtained from (7) comparable to (4) are

$$n_x = \frac{4/h_x - 1}{[R(I_X, y)/r]^{-2} - 1}. \qquad (9)$$

Hence values of n_x can be obtained from Table 2 using h_x for h and, for a pre-assigned value of $R(I_X, y)$ using the table with $R(I_X, y)/r$ in place of $R(I_Y, y)$. For example, with $r = 0.8$ and $h_x = 0.7$, in order to achieve a pre-assigned value of $R(I_X, y) = 0.56$, enter Table 2 with $0.56/0.8 = 0.7$ for $R(I_Y, y)$ and $h = 0.7$ and on doing so find that 5 progeny are needed. Note from (9) that the pre-assigned value of $R(I_X, y)$ cannot exceed the genetic correlation – as is eminently reasonable.

Similar relationships exist between (8) and Tables 3 and 4. From (3) and (8) we see that (8) is

$$RSE(I_X, Y, y) = pRSE(I_X, X, x),$$

so that values of (8) come from Table 3 by multiplication by p. Similarly, values of n_x obtained from (8) comparable to (5) are

$$n_x = \frac{4 - h_x}{(RSE^*)^{-2} - h_x} \qquad (10)$$

where

$$RSE^* = RSE(I_X, Y, y)/p \ . \qquad (11)$$

Thus values of n_x can be obtained from Table 4 using h_x for h and, for a pre-assigned value of $RSE(I_X, Y, y)$, using the table with RSE^* in place of RSE. For example, with $p = 0.8$ and $h_x = 0.4$, in order to achieve a pre-assigned value of $RSE(I_X, Y, y) = 1.24$ enter Table 4 with $RSE^* = 1.24/0.8 = 1.55$ for RSE and $h = 0.4$, and see that 222 progeny are needed. Notice from (10) that the upper limit on RSE^* is $h_x^{-\frac{1}{2}}$, and hence from (11) the corresponding upper limit on $RSE(I_X, Y, y)$ is $h_x^{-\frac{1}{2}}p = rh_y^{-\frac{1}{2}}$, using (6). Also, if this upper limit is to exceed unity then $r > \sqrt{h_y}$ must be satisfied.

3.2. Progeny-tests with the same numbers of progeny

We now consider the relative selection efficiency of I_X compared to I_Y. From (3)

$$RSE(I_Y, Y, y) = \sqrt{\frac{n_y}{4 + h_y(n_y - 1)}} \qquad (12)$$

and so, on defining

$$q = RSE(I_X, I_Y, y) = \frac{RSE(I_X, Y, y)}{RSE(I_Y, Y, y)}$$

in the manner of Searle [1965] we get, using (8),

$$q = p\sqrt{\frac{n_x[4 + h_y(n_y - 1)]}{n_y[4 + h_x(n_x - 1)]}} \quad . \tag{13}$$

This can be rearranged as

$$q = p\sqrt{\frac{n_x}{4 + h_x(n_x - 1)}} \Big/ \sqrt{\frac{n_y}{4 + h_y(n_y - 1)}} \tag{14}$$

where each square root term is an RSE like (12) and its value can be obtained from Table 3. A study of the behavior of q for variations in r, h_x, h_y, n_x and n_y is difficult. However, some conclusions can be reached, particularly on assuming the same number of progeny whether using X or Y. This is by no means an impractical assumption and so to begin with we take

$$n_x = n_y = n \tag{15}$$

and have

$$q = p\sqrt{\frac{4 + h_y(n - 1)}{4 + h_x(n - 1)}} = r\sqrt{\frac{4h_x + h_xh_y(n - 1)}{4h_y + h_xh_y(n - 1)}} \quad . \tag{16}$$

Progeny-testing with n progeny will be better using the alternative trait X than using the basic trait Y when $q > 1$. (We use 'better' and 'poorer' in the sense of bringing faster – and slower – improvement in y.) From (16) it is clear that $q > 1$ occurs when

$$p > \sqrt{\frac{4 + h_x(n - 1)}{4 + h_y(n - 1)}} \tag{17}$$

and, equivalently,

$$r > \sqrt{\frac{4h_y + h_xh_y(n - 1)}{4h_x + h_xh_y(n - 1)}} \quad . \tag{18}$$

Because, by definition, $r \leq 1$, (18) can be true only for $h_y < h_x$, in which case the right-hand side of (17) exceeds unity, i.e., $p > 1$. Furthermore, as is evident in (6), $p > 1$ only when $r > \sqrt{h_y/h_x}$ and one can readily show that the right-hand side of (18) always exceeds $\sqrt{h_y/h_x}$. Hence whenever (18) is satisfied so is $r > \sqrt{h_y/h_x}$ and so both $q > 1$ and $p > 1$; moreover, q can never exceed unity without p doing so also. However, for $h_y < h_x$ the first expression of (16) shows that $q < p$. Thus we have the following situation: progeny-testing with n progeny will be better using X rather than Y (i.e., $q > 1$) only when $h_y < h_x$ and when (18) is satisfied; and then $p > q > 1$, implying that the increased rate of improvement in y will be greater under mass selection than under progeny-testing. Table 5 gives

(SHOW TABLE 5)

examples of the lower limit of r imposed by (18) for $h_x > h_y$ and for n = 10, 30, 50 and 100.

Of particular importance in all this is that only if the alternative trait has higher heritability than the basic trait, can it be beneficial to use the alternative trait; and it will be beneficial only if the genetic correlation satisfies (18), e.g., exceeds the values shown in Table 5. Furthermore, these lower limits approach 1.00 as $n \to \infty$. This means that only for small numbers of progeny is there much of a range of values of r that permits $q > 1$ to be satisfied. This is because $q \to r$ as $n \to \infty$, as is evident from (16).

Table 5 is based on expressing the inequality $q > 1$ with $n_x = n_y = n$ in terms of limits on r. But it can also be expressed in terms of n as

$$n < L_1 \quad \text{for} \quad L_1 \equiv L_1(h_x, h_y, r) = 1 + \frac{4(r^2 h_x - h_y)}{h_x h_y (1 - r^2)} \tag{19}$$

$$= 1 + \frac{4(p^2 - 1)}{h_x (1 - r^2)} . \tag{20}$$

Clearly L_1 is positive for $p > 1$, so that when (17) and (18) are satisfied so also is $n < L_1$. Thus is established an upper limit L_1 such that progeny-testing using an alternative trait is better than using the basic trait with the same number of progeny. This means that for a particular alternative trait with $p > 1$ progeny-testing based on that trait can only be better than using the basic trait with the same number of progeny up to a limited number of progeny – and beyond that limit progeny-testing using the basic trait is better. Table 6 shows values of this limit for situations in which it exists, namely for the conditions under which $p > 1$, which are $r > \sqrt{h_y}$ and $h_x > h_y/r^2$, as discussed in Searle [1965]. For

(SHOW TABLE 6)

example, when $h_y = 0.2$ and $r = 0.9$ $p > 1$ if $h_x > 0.25$, and in fact if $h_x = 0.6$ then $p = 0.9\sqrt{0.6/0.2} = 1.56$; and from Table 6 we find that in this case only for 51 progeny or less, will progeny-testing using the alternative trait be better than using the basic trait, when both progeny-tests have the same number of progeny. For more than 51 progeny, using the basic trait will be better.

3.3. Progeny-tests with different numbers of progeny

We now consider progeny-tests based on different numbers of progeny: n_x using the alternative trait and n_y using the basic trait – as in equation (14). We have just seen when $p > 1$ that for n below the limits L_1 given in Table 6, progeny-testing with $n_x = n_y = n$ progeny is better using the alternative trait than using the basic trait, i.e., $q > 1$. For some cases of $n_x = n_y < L_1$, the value of q will exceed 1.00 sufficiently as to enable use of $n_x < n_y$ and still have $q > 1$. This means that for some values of $n_y < L_1$ using fewer progeny than n_y on the alternative trait (i.e., $n_x < n_y$) will be better than using n_y progeny with the basic trait. For example, from Table 6, with $r = 0.9$, $h_y = 0.2$ and $h_x = 0.6$, we saw

that $L_1 = 51$; and for $n_x = n_y = n = 20 < L_1 = 51$, equation (16) gives $q = 1.11$. This exceeds 1.00 sufficiently that we can in fact use the alternative trait with $n_x < 20$ and from (14) still have $q > 1$, i.e., still have the alternative trait being better than the basic trait. Thus equation (14) gives $q = 1.07$ for $n_x = 15$, and $q = 1.04$ for $n_x = 12$, so that either 15 or 12 progeny using the alternative trait are better than 20 progeny using the basic trait. Clearly, 15 or 12 progeny on the alternative trait ($q = 1.07$ and $q = 1.04$, respectively) are not as good as 20 progeny are ($q = 1.11$), but the important thing is that in these cases q still exceeds unity but with $n_x < n_y$.

Relationships between n_x and n_y can be considered in more general terms through the use of (14). Suppose we ask "how many progeny are needed using X to be equivalent to progeny-testing with n_y progeny using Y?" The answer is the solution for n_x to the equation $q = 1$. Using (14), this is

$$n_x = n_x(n_y, h_x, h_y, r) = \frac{n_y h_y(4 - h_x)}{h_x[r^2(4 - h_y) - n_y h_y(1 - r^2)]} . \qquad (21)$$

Since $n_x > 0$ we immediately see that this equivalence can occur only if

$$n_y < L_2 \quad \text{for} \quad L_2 \equiv L_2(h_y, r) = \frac{r^2(4 - h_y)}{h_y(1 - r^2)} . \qquad (22)$$

Since we have already seen that this kind of equivalence occurs for $n_x = n_y = n < L_1$ of (19), we would expect L_2 of (22) to exceed L_1 of (19); and it does, as may be easily shown. Thus L_2 of (22) is an upper limit to the number of progeny using the basic trait, above which not even an infinite number of progeny using the alternative trait can yield equivalent progeny-testing procedures. For example, with $r = 0.9$, $h_y = 0.2$, the limit given by (22) is $L_2 = 81$. Hence progeny-testing

using alternative traits with genetic correlation 0.9 can never be preferred to using the basic trait with 81 or more progeny, no matter how many progeny are available for the alternative trait, nor what the heritability of that alternative trait may be. For example, with $n_y = 100 > L_2 = 81$ and $n_x = 200$, with $h_x = 0.6$, equation (14) yields $q = 0.97 < 1$. Thus L_1 of (19) and L_2 of (22) divide values of n_y into 3 groups:

Group I: for $n_y < L_1$, there exists $n_x < n_y$ such that using the alternative trait is preferred.

Group II: for $L_1 \leq n_y < L_2$, there exists $n_x > n_y > L_1$ such that using the alternative trait is preferred.

Group III: for $n_y > L_2$, using the alternative trait can never be preferred.

The example of $h_x = 0.2$, $h_y = 0.6$ and $r = 0.9$ that has been used throughout this section is illustrated in Figure 1, which shows $R(I_Y, y)$ and $R(I_X, y)$ plotted against n. The division of n-values into 3 groups, insofar as n_y is concerned, is self-evident.

(SHOW FIGURE 1)

The division between Group I and Group II values of n_y is L_1, as discussed in the preceding section and tabulated in Table 6. The division between Groups II and III is L_2 of (22). For Group II values of n_y, where $n_x > n_y$, the lack of complete correlation between the alternative and basic traits (i.e., $r < 1$) can be compensated for in using the alternative trait by having more progeny, but only to a certain extent. There is a definite limit beyond which this compensatory effect cannot be had, and that limit is L_2 which represents the number of progeny using the basic trait beyond which not any number of progeny using the alternative trait

can be equivalent. Values of L_2 are shown in Table 7, for r from 0.50 to 0.98 and

(SHOW TABLE 7)

for h_y from 0.1 to 1.0. Tabulation of L_2 for $r < 0.50$ is not given because L_2 is then less than 10 (for $h_y > 0.10$). As is evident in Table 7 and from (22), values of L_2 are smaller for small values of r than for large, and also for large values of h_y than for small. In particular, when the basic trait has a heritability of 0.60 or more, L_2 is less than 20 unless r is 0.90 or more. This simply means that in these situations using an alternative trait can only be equivalent to using a basic trait having high heritability when the latter is used on relatively few progeny. When the progeny-test using the basic trait has more than 20 or so progeny it is then more reliable than any progeny-test using an alternative trait could be, no matter how many progeny were available. On the other hand, L_2 is large when r is close to 1.00, corresponding to the upper limit on n_y being infinite when r = 1. In this case the high genetic correlation between the alternative and basic traits means that the two traits are sufficiently correlated genetically that a progeny-test using the alternative trait can be equivalent to one using the basic trait even for large numbers of progeny in the latter. When they are perfectly correlated, r = 1, the limit is infinite, meaning that for all progeny-tests using the basic trait equivalent tests using the alternative trait can be found, and in that case $n_x = n_y(4/h_x - 1)/(4/h_y - 1)$. A final and obvious comment, but one worthy of note, is that L_2 is independent of h_x, the heritability of the alternative trait.

With n_x of (21) being a function of 4 variables, there is difficulty in presenting a table of values that is of both reasonable size and practical use. Table 8 is but a small example, for n_y = 10, 15, 20 and 25, for h_y = 0.3, 0.6 and 0.5, and for r = 0.75 to 0.95. There are no entries for combinations of these

(SHOW TABLE 8)

values for which n_y exceeds its limit (see Table 7). Interpolation for intermediate cases is not recommended due to the nature of the function, and direct calculation from the formula should be made for cases not shown. Generally speaking, for a given value of n_y, the value of n_x increases with increases in r and h_y; and it decreases with increases in h_x. Such results are as one would expect and are demonstrated by the values in Table 8.

4. Using Two Alternative Traits

Suppose we wish to compare progeny-tests using two different alternative traits X_1 and X_2 for selection for improvement in y. Let h_1, h_2 be their heritabilities, r_1, r_2 their genetic correlations with the basic trait and n_1, n_2 the number of progeny used. Then the relative selection efficiency of I_{X_1} compared to I_{X_2} is

$$RSE(I_{X_1}, I_{X_2}, y) = R(I_{X_1}, y)/R(I_{X_2}, y)$$

$$= \frac{r_1}{r_2}\sqrt{\frac{n_1h_1}{4 + h_1(n_1 - 1)}} \Big/ \frac{n_2h_2}{4 + h_2(n_2 - 1)}, \qquad (23)$$

using (7). Values of the square root terms in (23) can be obtained from Table 1.

Progeny-testing with n_1 progeny using an alternative trait X_1 will be equivalent to having n_2 progeny using another alternative X_2 when $R(I_{X_1}, I_{X_2}, y) = 1$. Using (23) and solving for n_1 gives

$$n_1 = \frac{n_2r_2^2h_2(4 - h_1)}{h_1[r_1^2(4 - h_2) - n_2h_2(r_2^2 - r_1^2)]}, \qquad (24)$$

a result very similar to n_x of (21). In fact, if we let

$$\rho_1 = r_1^2/r_2^2 \quad \text{for} \quad r_1^2 < r_2^2 \tag{25}$$

then (24) can be expressed as a form of (21) in the manner $n_1 = n_x(n_2, h_1, h_2, \rho_1)$.

Consideration of when progeny-testing using X_1 will be better than using X_2 (in the sense of faster improvement in y) is achieved through studying

$$n_1 - n_2 = \frac{n_2}{h_1}\left[\frac{4(r_2^2h_2 - r_1^2h_1) + (n_2 - 1)h_1h_2(r_2^2 - r_1^2)}{r_1^2(4 - h_2) - n_2h_2(r_2^2 - r_1^2)}\right], \tag{26}$$

available from (24). Because the denominator of (23) is always positive when $r_1^2 > r_2^2$, a situation that cannot occur with (21), we need to distinguish three cases, according as $r_1^2 <, =,$ or $> r_2^2$. In all cases, manipulation of (24) and (26) leads to the results that follow.

Case 1, $r_1^2 < r_2^2$, $\rho_1^2 = r_1^2/r_2^2 < 1$

When $n_2 < L_2(h_2, \rho_1)$, X_1 is better than X_2 with $n_1 > n_2$ except that when $h_1/h_2 > 1/\rho_1^2$ and $n_2 < L_1(h_1, h_2, \rho_1) < L_2(h_2, \rho_1)$, then X_1 can be better with $n_1 < n_2$. And when $n_2 \geq L_2(h_2, \rho_1)$ the alternative X_1 can never be better than X_2.

Case 2, $r_1^2 = r_2^2$

Equation (24) reduces to

$$n_1 = \lambda n_2 \quad \text{for} \quad \lambda = \frac{4/h_1 - 1}{4/h_2 - 1}. \tag{27}$$

It is always possible to have X_1 better than X_2 for $n_1 > n_2$, and for $h_1 > h_2$ we can have X_1 better than X_2 for $n_1 < n_2$.

Case 3, $r_1^2 > r_2^2$, $\rho_2^2 = r_2^2/r_1^2 < 1$

X_1 can always be better than X_2 with $n_1 < n_2$ when $h_1/h_2 \geq \rho_2^2$; and for $h_1/h_2 < \rho_2^2$, X_1 can be better than X_2 with $n_1 < n_2$ when $n_2 \geq L_1(h_2, h_1, \rho_2)$ but only with $n_1 > n_2$ when $n_2 < L_1(h_2, h_1, \rho_2)$.

Additional results similar to these could be derived for combining the alternative traits into selection indices either with each other or with the basic trait, just as was done in the case of mass selection in Searle [1965]. However, the algebraic manipulations increase in complexity and it becomes difficult to establish any usable conditions under which some alternatives are preferable to others.

5. Sampling Variances

Discussion of whether an alternative trait is to be preferred or not has been entirely in terms of genetic parameters. In practice these have to be estimated. Decisions about using an alternative trait therefore have to be made on the basis of estimated RSE's, for example, estimated values of q of (14). Even though the sampling distribution of these estimates is unknown, there would be some (slight) comfort to using the estimates if we could have even approximate standard errors for them. This has been done in the case of mass selection, where p of (6) is the appropriate measure of relative selection efficiency. Approximate sampling variances of an estimator of p have been obtained by Searle [1965], based on parent-progeny records for estimating genetic parameters and by Scheinberg [1967] based on sibship data. In both cases derivation of the sampling variance is tedious, is approximate only, and results in an expression that is too complicated to be studied analytically. Only numerical studies are possible. These difficulties are even worse in the case of the sampling variance of the relative

selection efficiency of progeny-testing, of say $RSE(I_Y, Y, y)$ of (3) or of q of (14). The added difficulty compared to dealing with p is that whereas p is just a simple product of r, $h_x^{\frac{1}{2}}$ and $h_y^{-\frac{1}{2}}$, q involves $[4 + h_x(n_x - 1)]^{-\frac{1}{2}}$ and $[4 + h_y(n_y - 1)]^{\frac{1}{2}}$; and the methods of deriving sampling variances do not lead themselves at all easily to handling this kind of expression. For example, adapting Scheinberg's [1967] notation slightly, his expression (8) is

$$\hat{p} = \hat{G}_{xy}\hat{P}_{yy}^{\frac{1}{2}}\hat{P}_{xx}^{-\frac{1}{2}}\hat{G}_{yy}^{-1}$$

where $\hat{P}_{yy}$ (and $\hat{G}_{yy}$) are estimated phenotypic (and genotypic) variances – and $\hat{G}_{xy}$ is an estimated genotypic covariance. The comparable expression for $\hat{q}$ is

$$\hat{q} = \left(\sqrt{n_x/n_y}\right)\hat{G}_{xy}\hat{H}_{yy}^{\frac{1}{2}}\hat{H}_{xx}^{-\frac{1}{2}}\hat{G}_{yy}^{-1}$$

for

$$H_{yy} = 4P_{yy} + (n_y - 1)G_{yy} .$$

The complications that result in trying to follow Scheinberg's procedure are horrendous – and even if followed successfully yield expressions considerably more complicated than his. Furthermore, they have the added complication of involving n_x and n_y. Their practical value would therefore appear to be very limited.

A feasible alternative might be to obtain

$$\mathrm{var}(\hat{q}) \doteq (q/p)^2\mathrm{var}(\hat{p})$$

directly from (14). Cautious use of this in the form

$$\widehat{\mathrm{var}}(\hat{q}) \doteq (\hat{q}/\hat{p})^2\widehat{\mathrm{var}}(\hat{p})$$

might be little worse than the excessively complicated procedure of the preceding

paragraph. Calculation of $\widehat{var}(\hat{p})$ would then be derived from Scheinberg [1967], which, in its use of sib records, is the appropriate form to use here for progeny-testing.

6. Conclusions

Space precludes summarizing all the conclusions that could be drawn from the preceding discussion but some of the more important ones are now listed. In all cases these conclusions pertain to progeny-testing.

1. The relative selection efficiency of progeny-testing compared to performance testing is

$$RSE(I_Y, Y, y) = \sqrt{n/[4 + h(n - 1)]} \ ,$$

with a limiting value of $h^{\frac{1}{2}}$ as $n \to \infty$. [See equation (3) and Table 3.]

2. The number of progeny needed so that $RSE(I_Y, Y, y)$ dquals some pre-assigned value RSE is

$$n = (RSE)^2(4/h - 1)/[1/h - (RSE)^2] \ .$$

[See equation (5) and Table 4.]

3. With $p = r\sqrt{h_x/h_y}$ being the RSE of the alternative trait X under mass selection [equation (6)], the RSE of progeny-testing with X compared to performance testing with Y is

$$RSE(I_X, Y, y) = p\sqrt{n_x/[4 + h_x(n_x - 1)]} \ .$$

Values can be calculated from Table 3 by multiplying by p. [See equation (8).]

4. The number of progeny needed so that $RSE(I_X, Y, y)$ equals some pre-assigned value RSE is

$$n_x = (RSE^*)^2(4/h_x - 1)/[1/h_x - (RSE^*)^2]$$

where $RSE^{*} = RSE/p$. [See equations (10) and (11) and Table 4.]

5. The RSE of progeny-testing with n_x progeny using the alternative trait X compared to n_y progeny with the basic trait Y is

$$q = RSE(I_X, I_Y, y)$$

$$= p\sqrt{\frac{n_x/[4 + h_x(n_x - 1)]}{n_y/[4 + h_y(n_y - 1)]}} = p\,\frac{RSE(I_X, X, x)}{RSE(I_Y, Y, y)} .$$

[See equation (14).]

6. Progeny-testing with n progeny will be better using X rather than Y (i.e., $q > 1$) only when

$$h_x > h_y \quad \text{and} \quad r > \sqrt{\frac{4h_y + h_x h_y(n - 1)}{4h_x + h_x h_y(n - 1)}} ,$$

in which case $p > q > 1$. [See equation (18) and Table 5.] Equivalent conditions are

$$h_x > h_y \quad \text{and} \quad n_y < L_1$$

for

$$L_1 = L_1(h_x, h_y, r) = 1 + \frac{4(r^2 h_x - h_y)}{h_x h_y(1 - r^2)} .$$

[See equation (19) and Table 6.]

7. Progeny-testing with n_y progeny using Y is equivalent ($q = 1$) to using n_x progeny with X for

$$n_x = \frac{n_y h_y(4 - h_x)}{h_x[r^2(4 - h_y) - n_y h_y(1 - r^2)]}$$

provided

$$n_y < L_2 \quad \text{for} \quad L_2 = L_2(h_y, r) = \frac{r^2(4 - h_y)}{h_y(1 - r^2)} .$$

[See equation (22) and Tables 7 and 8.]

8. L_1 and L_2 divide values of n_y into 3 groups:

   (1) $n_y < L_1$ where X can be preferred for $n_x < n_y$;

   (2) $L_1 \leq n_y < L_2$ where X can be preferred for $n_x > n_y$;

   (3) $L_2 \leq n_y$ where X can never be preferred.

[See section 3.3 and Figure 1.]

9. The RSE of progeny-testing with n_1 progeny on X_1 compared to n_2 on X_2 is

$$RSE(I_{X_1}, I_{X_2}, y) = \frac{r_1}{r_2}\sqrt{\frac{n_1 h_1/[4 + h_1(n_1 - 1)]}{n_2 h_2/[4 + h_2(n_2 - 1)]}} .$$

[See equation (23).] Situations where X_1 is preferred to X_2 are discussed in Section 4.

References

[1] Ch'ang, T. S. and Rae, A. L. [1972]. The genetic basis of growth, reproduction and maternal environment in Romney ewes. II. Genetic covariance between hogget characters, fertility and maternal environment of the ewe. Australian J. of Agric. Research 23, 149-165.

[2] Eisen, E. J. [1966]. Effect of the biometrical relationship among total milk yield, milk constituent and per cent of milk constituent on response to selection. J. Dairy Science 49, 1230-1234.

[3] Falconer, D. S. [1960]. Introduction to Quantitative Genetics. Oliver and Boyd, Edinburgh.

[4] Festing, M. and Nordskog, A. W. [1967]. Response to selection for body weight and egg weight in chickens. Genetics 55, 219-231.

[5] Nordskog, A. W., Festing, M. and Verghese, M. W. [1967]. Selection for egg production and correlated responses in the fowl. Genetics 55, 179-191.

[6] Scheinberg, E. [1967]. The sampling variance of the relative efficiency of indirect to direct selection when using variance-covariance components. Australian J. of Statistics 9, 35-40.

[7] Searle, S. R. [1963]. The efficiency of ancestor records in animal selection. Heredity 18, 351-360.

[8] Searle, S. R. [1964]. Progeny-tests of sire and son. J. Dairy Science 47, 414-420.

[9] Searle, S. R. [1965]. The value of indirect selection: I. Mass selection. Biometrics 21, 682-707.

[10] Singh, R. K., Bellman, K. and Ahrens, H. [1967]. Die Abhängigheit der korrelierten Antwort auf die Selektion von Kopplung und Heritabilität. Biometrische Zeitschrift 9, 240-249.

[11] Thompson, N. R. and Loganathan, S. [1968]. Composition of cows' milk. II. Genetic influences. J. Dairy Science 51, 1933-1935.

[12] Wilton, J. W. and Van Vleck, L. D. [1968]. Selection of dairy cows for economic merit. J. Dairy Science 51, 1680-1688.

TABLE 1

Correlation* between progeny-test and additive genotype of a parent.

No. of progeny (n)	Heritability (h)									
	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
2	.22	.31	.37	.43	.47	.51	.55	.58	.61	.63
6	.36	.49	.57	.63	.68	.72	.75	.77	.80	.81
10	.45	.59	.67	.72	.77	.80	.82	.85	.86	.88
15	.53	.66	.74	.79	.82	.85	.87	.89	.90	.91
20	.59	.72	.79	.83	.86	.88	.90	.91	.92	.93
25	.62	.75	.82	.86	.88	.90	.92	.93	.94	.94
30	.66	.78	.84	.87	.90	.92	.93	.94	.95	.95
40	.72	.82	.87	.90	.92	.94	.95	.95	.96	.96
50	.75	.85	.89	.92	.94	.95	.95	.96	.97	.97
60	.78	.87	.91	.93	.95	.95	.96	.96	.97	.98
80	.82	.90	.93	.95	.96	.97	.97	.98	.98	.98
100	.85	.92	.94	.96	.97	.97	.98	.98	.98	.99
200	.91	.96	.97	.98	.98	.98	.98	.99	.99	.99

* Equation (2): $R(I_Y, y) = \sqrt{nh/[4 + (n - 1)h]}$.

TABLE 2

Number* of progeny needed in a progeny test in order to estimate the additive genotype of a parent with pre-assigned value of the correlation between the progeny-test and parent's genotype.

Pre-assigned correlation $R(I_Y, y)$	Heritability (h)									
	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
.20	2	1	1	1	1	1	1	1	1	1
.40	7	4	3	2	2	2	1	1	1	1
.50	13	7	5	3	3	2	2	2	2	1
.60	22	11	7	6	4	4	3	3	2	2
.70	38	18	12	9	7	6	5	4	4	3
.80	70	34	22	16	13	11	9	8	7	6
.90	167	81	53	39	30	25	21	18	15	13
.95	361	176	115	84	65	53	44	38	32	28
.98	946	461	300	219	170	138	115	98	84	73
.99	1,921	936	608	444	345	280	233	198	170	148

* Equation (4): $$n = \frac{4/h - 1}{[R(I_Y, y)]^{-2} - 1}.$$

TABLE 3

Relative selection efficiency* of a progeny-test on n progeny.

No. of progeny (n)	Heritability (h)									
	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
2	.70	.69	.68	.67	.67	.66	.65	.65	.64	.63
6	1.15	1.10	1.04	1.00	.96	.93	.89	.86	.84	.81
10	1.40	1.31	1.22	1.15	1.08	1.03	.99	.94	.91	.88
15	1.67	1.49	1.35	1.25	1.17	1.10	1.04	.99	.95	.91
20	1.84	1.60	1.44	1.31	1.22	1.14	1.08	1.02	.97	.93
25	1.98	1.69	1.49	1.36	1.25	1.17	1.10	1.04	.99	.94
30	2.09	1.75	1.54	1.39	1.27	1.18	1.11	1.05	1.00	.95
40	2.25	1.84	1.60	1.43	1.30	1.21	1.13	1.06	1.01	.96
50	2.37	1.90	1.64	1.46	1.32	1.22	1.14	1.08	1.02	.97
60	2.46	1.95	1.66	1.47	1.34	1.23	1.15	1.08	1.03	.98
80	2.59	2.01	1.70	1.50	1.36	1.25	1.16	1.09	1.03	.98
100	2.68	2.05	1.73	1.51	1.37	1.26	1.17	1.10	1.04	.99
200	2.89	2.14	1.77	1.55	1.39	1.27	1.18	1.11	1.05	.99
∞	3.16	2.24	1.83	1.58	1.41	1.29	1.20	1.12	1.05	1.00

* Equation (3): $RSE(I_Y, Y, y) = \sqrt{n/[4 + h(n - 1)]}$.

TABLE 4

Number* of progeny needed so that progeny-testing is at least as good as performance testing. (Upper limits on RSE shown in parentheses.)

	Heritability (h)									
RSE	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
1.00	5	5	6	6	7	9	11	16	31	(1.00)
1.05	5	6	7	8	9	11	16	30	441	
1.10	6	7	8	9	11	16	27	121	(1.054)	
1.15	6	7	9	11	14	22	59	(1.12)		
1.20	7	8	10	13	18	36	(1.195)			
1.25	8	9	11	15	25	85				
1.40	10	13	18	33	343	(1.29)				
1.50	12	16	26	81	(1.41)					
1.55	13	18	32	222						
1.60	14	20	41	(1.58)						
1.80	19	35	428							
2.00	26	76	(1.83)							
2.20	37	575								
2.50	65	(2.24)								
3.00	351									
3.10	961									
	(3.16)									

* Equation (5): $n = \dfrac{4 - h}{(RSE)^{-2} - h}$.

TABLE 5

Lower limit* on r, the genetic correlation between an alternative trait and a basic trait, such that a progeny-test using n progeny is better based on the alternative trait than the basic trait.

Heritability of basic trait (h_y)	Heritability of alternative trait (h_x)											
	.4				.6				.8			
	No. of progeny (n)				No. of progeny (n)				No. of progeny (n)			
	10	30	50	100	10	30	50	100	10	30	50	100
.2	.81	.89	.93	.95	.73	.85	.90	.94	.69	.83	.89	.93
.4					.91	.96	.97	.98	.85	.93	.95	.98
.6									.94	.95	.96	.96

* Equation (18): $r > \sqrt{\dfrac{4h_y + h_x h_y (n-1)}{4h_x + h_x h_y (n-1)}}$, with $h_x > h_y$.

TABLE 6

Maximum number* of progeny using the basic trait, for which progeny-testing using an alternative trait can be equal or better when using less progeny.

Heritability of basic trait (h_y)	Heritability of alternative trait (h_x)								
	.2	.3	.4	.5	.6	.7	.8	.9	1.0
				Genetic correlation = 0.4					
.1						1	2	3	3
				Genetic correlation = 0.5					
.1			1	3	5	6	7	8	9
.2							1	1	2
				Genetic correlation = 0.6					
.1		2	7	11	13	14	15	16	17
.2					1	3	4	5	6
.3								1	2
				Genetic correlation = 0.7					
.1		13	19	23	26	28	29	30	31
.2				4	7	9	10	11	12
.3						2	4	5	5
.4									2
				Genetic correlation = 0.8					
.1	16	35	44	49	53	56	58	59	61
.2			8	14	18	20	22	24	25
.3				2	6	8	10	12	13
.4						2	4	6	7
.5							1	2	4
				Genetic correlation = 0.9					
.1	66	101	118	129	136	141	145	148	150
.2		16	33	44	51	56	59	62	65
.3			5	16	23	28	31	34	36
.4				1	8	13	17	20	22
.5						5	8	11	14
.6							3	6	8
.7								1	4
.8									1

* Equation (19): $\text{Maximum} = L_1 = 1 + \dfrac{4(r^2 h_x - h_y)}{h_x h_y (1 - r^2)}$.

TABLE 7

Maximum number* of progeny using the basic trait for which progeny-testing using an alternative trait can be equivalent or better.

Heritability of basic trait (h_y)	Genetic correlation, r, between basic and alternative traits								
	.50	.60	.70	.75	.80	.85	.90	.95	.98
.1	13	21	37	50	69	101	166	360	946
.2	6	10	18	24	33	49	81	175	460
.3	4	6	11	15	21	32	52	114	299
.4	3	5	8	11	16	23	38	83	218
.5	2	3	6	9	12	18	29	64	169
.6	1	3	5	7	10	14	24	52	137
.7	1	2	4	6	8	12	20	43	114
.8	1	2	3	5	7	10	17	37	97
.9	1	1	3	4	6	8	14	31	83
1.0	1	1	2	3	5	7	12	27	72

* Equations (21) and (22): Maximum = $L_2 = \dfrac{r^2(4 - h_y)}{h_y(1 - r^2)}$.

TABLE 8

Number* of progeny required in a progeny-test using an alternative trait to be equivalent to a progeny-test using the basic trait.

Number of progeny in progeny-test using basic trait (n_y)	Heritability of basic trait (h_y)	Heritability of alternative trait (h_x)														
		.3					.4					.5				
		Genetic correlation between traits (r)					Genetic correlation between traits (r)					Genetic correlation between traits (r)				
		.75	.80	.85	.90	.95	.75	.80	.85	.90	.95	.75	.80	.85	.90	.95
10	.3	49	29	21	16	13	36	21	15	12	9	28	17	12	9	7
	.4	180	58	34	23	18	131	42	25	17	13	102	33	19	13	10
	.5		141	55	33	24		103	40	24	17		80	31	19	14
15	.3	494	75	39	26	20	361	55	29	19	14	281	43	23	15	11
	.4		514	80	42	28		376	58	31	21		292	45	24	16
	.5			207	66	39			151	48	28			118	38	22
20	.3		356	74	40	27		260	54	30	20		202	42	23	16
	.4			259	71	40			189	52	30			147	41	23
	.5				132	57				97	42				75	33
25	.3			157	59	36			115	43	26			89	34	21
	.4				122	55				89	40				69	31
	.5				336	80				245	58				191	46

* Equation (20): $$n_x = \frac{n_y h_y (4 - h_x)}{h_x [r^2(4 - h_y) - n_y h_y (1 - r^2)]} .$$

Correlations: $R(I_Y, y) = \sqrt{\frac{nh_y}{4 + (n - 1)h_y}}$ and $R(I_X, y) = r\sqrt{\frac{nh_x}{4 + (n - 1)h_x}}$

for $h_y = 0.2$, $h_x = 0.6$, and $r = 0.9$.

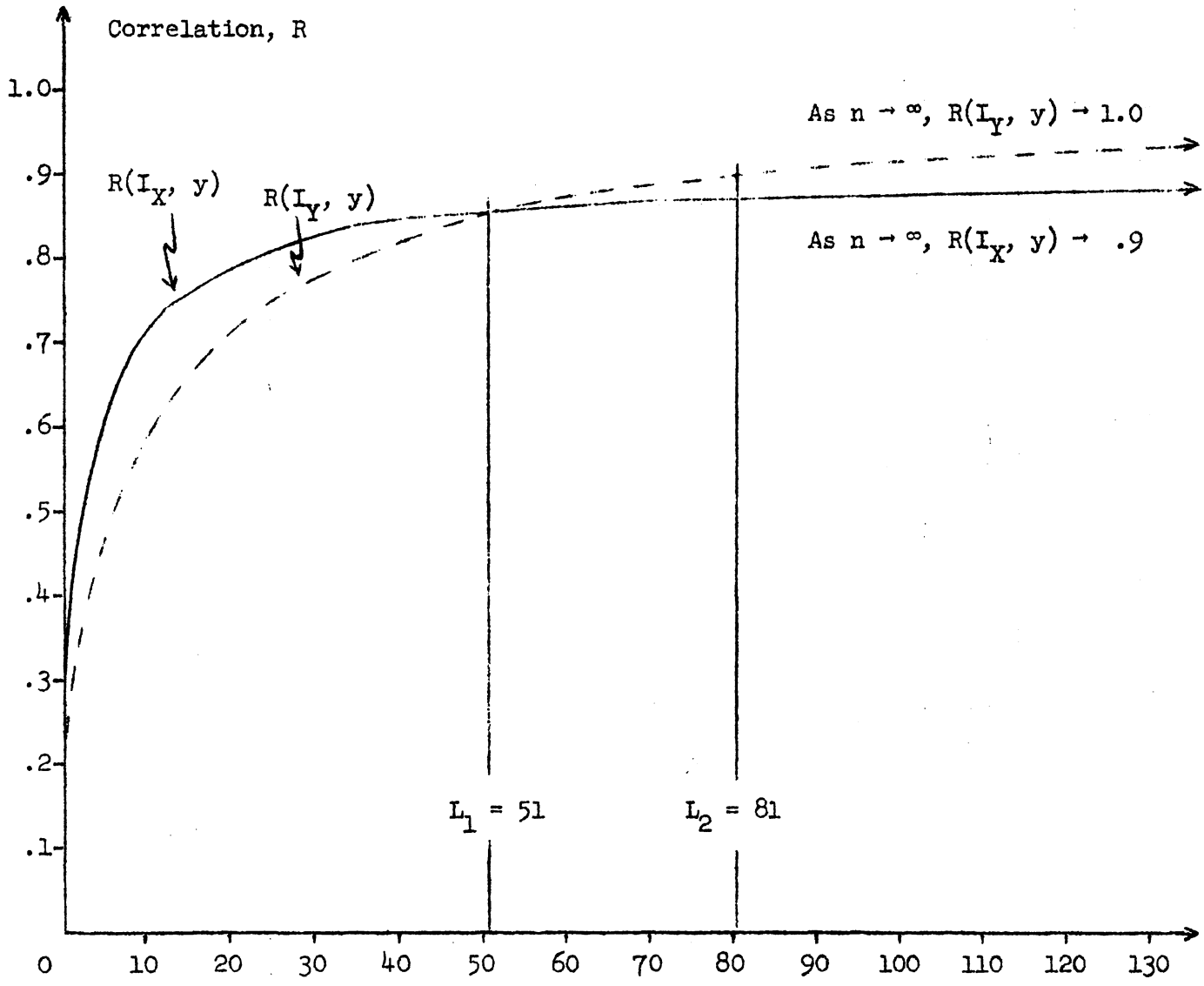

FIGURE 1

Appendix

Derivation of R_{Iy}, for $I = b\bar{Y}$, on n records.

$$b\sigma^2_{\bar{Y}} = \sigma_{\bar{Y}y} \Rightarrow b = \frac{\sigma_{\bar{Y}y}}{\sigma^2_{\bar{Y}}} = \frac{\frac{1}{2}\sigma^2_y}{\frac{n + \frac{1}{4}n(n-1)h}{n^2}\sigma^2_Y} = \frac{2nh}{4 + (n-1)h}$$

$$R_{Iy} = \frac{\sigma_{\bar{Y}y}}{\sigma_{\bar{Y}}\sigma_y} = \frac{\sigma_{\bar{Y}y}}{\sigma^2_{\bar{Y}}}\frac{\sigma_{\bar{Y}}}{\sigma_y} = \frac{2nh}{4 + (n-1)h}\sqrt{\frac{n + \frac{n}{4}(n-1)h}{n^2}}\frac{1}{h}$$

$$= \frac{2nh}{4 + (n-1)h}\sqrt{\frac{4 + (n-1)h}{4nh}} = \sqrt{\frac{nh}{4 + (n-1)h}}\,.$$

Derivation of $R(I_X, y)$

$$I = b\bar{X}, \quad b = \frac{\sigma_{\bar{X}y}}{\sigma^2_{\bar{X}}} = \frac{\frac{1}{2}r\sigma_x\sigma_y}{\frac{4 + (n-1)h_x}{4n}\sigma^2_X} = \frac{2rnh_x\sigma_y/\sigma_x}{4 + (n-1)h_x}$$

$$R(I_X, y) = \frac{\sigma_{\bar{X}y}}{\sigma_{\bar{X}}\sigma_y} = \frac{\sigma_{\bar{X}y}}{\sigma^2_{\bar{X}}}\frac{\sigma_{\bar{X}}}{\sigma_y}$$

$$= \frac{2rnh_x\sigma_y/\sigma_x}{4 + (n-1)h_x}\sqrt{\frac{4 + (n-1)h_x}{4n}}\frac{\sigma^2_X}{\sigma^2_y} = r\sqrt{\frac{nh_x}{4 + (n-1)h_x}}\,.$$